

# PASIG 2019

February 12-14, 2019 | El Colegio de  
México | Mexico City

## COMMUNITY LIGHTNING TALKS

**Mark Evans**, History Associates

**Danielle Calle**, MIAP

**Carl Wilson**, Open Preservation Foundation

**Sarah Romkey**, Artefactual Systems

**Amy Rudersdorf**, AVP

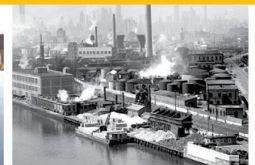
**#PASIG2019**

# HISTORY ASSOCIATES

## Preservation Case studies

Mark Evans – Director of Archives and Collection Management Services

PASIG 2019 – Mexico City





# History Associates: Capabilities



Tell Stories			Answer Questions	Understand, Preserve, Use	
Histories	Exhibits	Interpretive Planning	Historical Research	Archival Services	Collections Management
Cultural Heritage	Corporate	Government	Academic	Non-Profit	Private / Personal



# Case Study: Web Archiving



U.S. National Library of Medicine

## Project

Development and Implementation of full lifecycle web archiving workflow

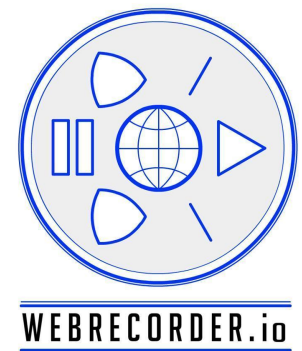
## Catalyst

Existing web collections in Archive-IT were incomplete, missing significant content and lacked metadata

## Solution

Development of workflow to include.

- Seed appraisal and selection
- Use of multiple crawl tools
- Test crawls to determine scope
- Quality Control steps and patch crawls
- Development of descriptive metadata



Example: Opioid epidemic - <https://archive-it.org/collections/9280>



# Case Study: Congressional Records

## Client

Various Senate and House offices

## Project

Office close out and preservation of member's legacy

## Challenges

- Short time frame for execution (2 months)
- Wide variety of record types and sources
- High degree of PII and
- Level of readiness of receiving repository
- **Extensive use of large scale proprietary systems with basic exit strategy (XML)**

## Solution ?

- Emulation not an option – restricted software license
- Data Normalization / What about functional behavior?
- Working group examining sustainability of an open source solution



■ simplicity®





# Come talk to us to learn more !!!

Mark Evans – [mevans@historyassociates.com](mailto:mevans@historyassociates.com)

Laura Starr – [lstarr@historyassociates.com](mailto:lstarr@historyassociates.com)

**Thank you!**

# Climate Change as a Prompt for the Digital Archive

*Investigando la materialidad de la preservación digital  
y el medio ambiente*

**Link: <https://bit.ly/2N681Sn>**

Danielle Calle, NYU Tisch  
Moving Image Archiving & Preservation  
[dcalle@nyu.edu](mailto:dcalle@nyu.edu)

---

# Towards a community-driven Archivematica

Sarah Romkey  
Archivematica Program Manager, Artefactual Systems  
PASIG 2019

---



# Archivematica

is a web- and standards-based, open-source application which allows your institution to preserve long-term access to trustworthy, authentic and reliable digital content.

---

# **Towards a community-driven Archivemania**

Why are moving in this direction?

- Better management of maintenance tasks through better planning
  - Alignment with our company philosophy
  - Desire to see as many people sustainably preserving digital heritage as possible
  - And to do that, we need to know we're working on the right things.
-

---

# Towards a community-driven Archivemania

What have we been up to?

- Implemented processes for filing issues
    - On Github
    - One repository to rule them all
    - Developed some guidelines for labels, etc
    - Implemented an “issue first” guideline.
  - Assigned a Product Owner
    - Part of Agile methodology
    - Responsible for triaging issues and gathering feedback from all stakeholders.
-

---

# Towards a community-driven Archivemania

What have we been up to?

- Implemented new tooling
    - Waffle: used to move Github issues from the “backlog” through to “done”: <https://waffle.io/artefactual/archivemania>
    - Trello: currently using for high-level release planning
    - Also using Trello as an interim road-mapping tool
-

---

# Towards a community-driven Archivemata

What are we doing next?

- Developing a product vision and roadmap
  - Implementing “request for change”/ADR [Architectural Decision Record] process
    - An initial proposal:  
<https://github.com/artefactual-labs/archivematica-rfcs-test>
  - Improve ability to work with external contributors
    - Laying out expectations more clearly
    - Community PRs part of triage process
    - MOU with Wellcome Collection to improve testing and contribute to each other's development efforts
-

---

# Community-driven community

Bi-monthly user forum:

[https://docs.google.com/document/d/1xowR7G9u\\_5HmJFLcnQRp5rWiPfocmjQ7o-2T6QVkJAu4/edit](https://docs.google.com/document/d/1xowR7G9u_5HmJFLcnQRp5rWiPfocmjQ7o-2T6QVkJAu4/edit)

Regional user groups in UK, Texas, USA Northeast,  
Netherlands/Flanders, Australasia

[https://wiki.archivematica.org/Community/Regional User Groups](https://wiki.archivematica.org/Community/Regional_User_Groups)

---



Carl Wilson, Technical Lead, OPF

JHOVE 2020

# What is JHOVE?

JSTOR/Harvard Object Validation Environment

An EXTENSIBLE software framework for format:

- identification,
- validation, and
- characterisation.





# Or Put Another Way...

JHOVE provides some answers for the following questions regarding a digital object:

- What format is it?
- The object is of format: F, is it?
- The object is of format: F, what are its properties?



# The Future of JHOVE

- Better process for module versioning
- Assign IDs for error messages
- Instrumentation/monitoring/measurement
- JHOVE webservice REST API
- Standardise underlying reporting mechanism
- Java version increment
- Updating JHOVE test corpus / error reporting coverage
- Strategies for user updates and major changes (pre major changes in 1.24)



JHOVE currently has the following modules:

Text formats : ASCII, XML, PDF, HTML

Image formats : GIF, JPEG, JPEG200, TIFF, PNG

Sound formats : WAVE, AIFF

Container formats : WARC, GZIP

The OPF also maintains veraPDF (PDF/A), jpylyzer (JPEG200), DPF Manager (TIFF)....



# JHOVE community

Use the GitHub Issue tracker: <https://github.com/openpreserve/jhove/issues>

- If there's something wrong.
- If you have an idea about something that can be improved.
- If you'd like to see a new feature.

You can also submit a contribution as a GitHub Pull Request:

<https://github.com/openpreserve/jhove/pulls>

- This can be a fix for the code base
- It could be improvements to the documentation or website
- It could even be a whole new module



# Help support open source software!

JHOVE is **open source**: it is free to download, use and modify.







However, it is not free of costs to **host, maintain, support** or **develop** the software. Creating **user guides** and **training materials** also requires resources. Currently, only OPF members support JHOVE financially as part of their membership fees.

OPF members get **technical support** to use our products and can participate in the Product Board to **prioritise development**

Join us: <https://openpreservation.org/about/join/>



# Resources

-  <http://jhove.openpreservation.org/> (documentation)
-  <http://lists.openpreservation.org/listinfo/jhove> (user mailing list)
-  <https://github.com/openpreserve/jhove> (code)
-  @openpreserve
-  [info@openpreservation.org](mailto:info@openpreservation.org)
-  <http://openpreservation.org/>

